Decision Making: who's in charge?

Nicolas P. Rougier

DECISION MAKING: WHO'S IN CHARGE?

NICOLAS P. ROUGIER

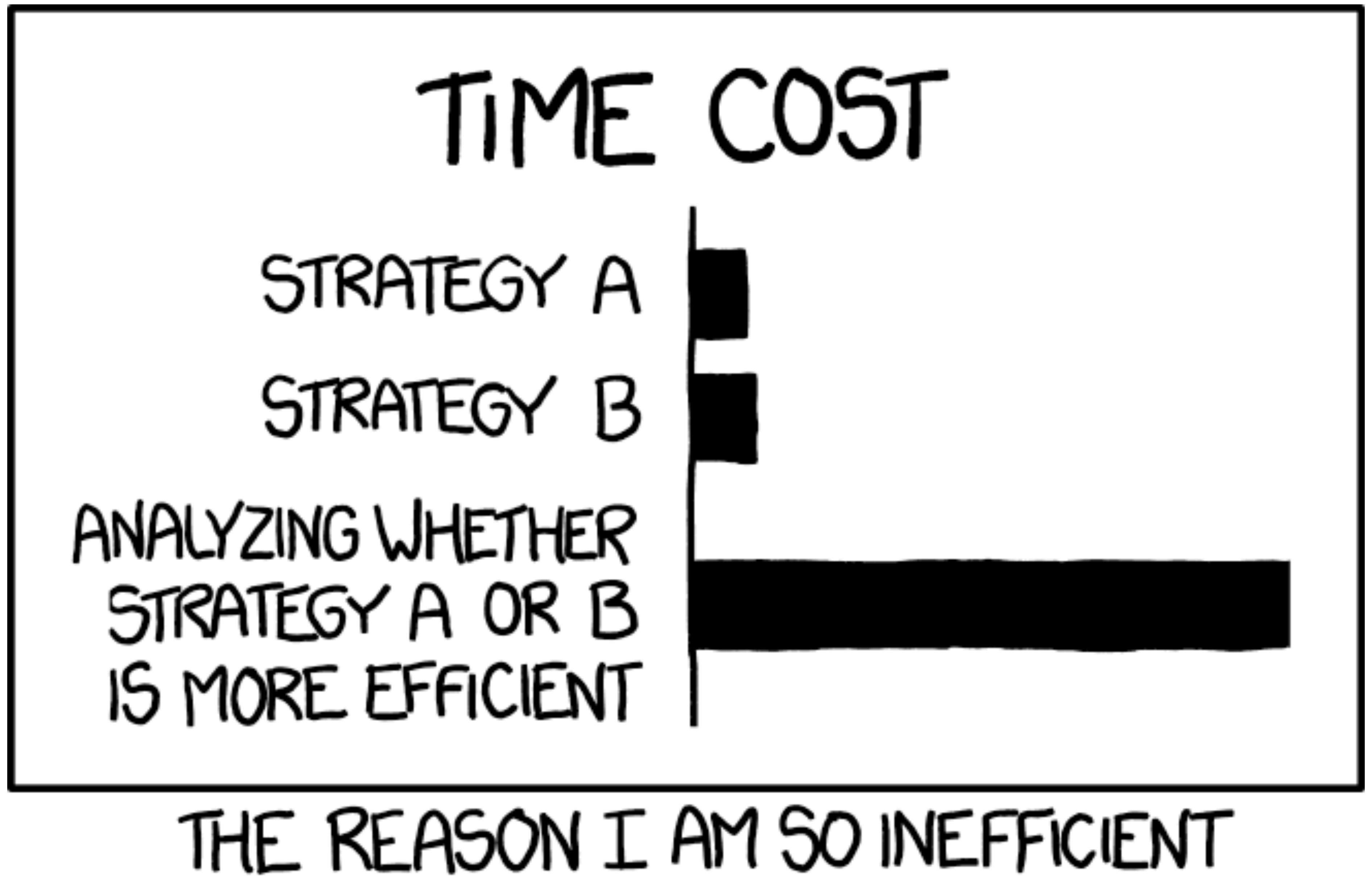
INRIA BORDEAUX /
INSTITUT DES MALADIES
NEURODEGENERATIVES

A TENTATIVE DEFINITION

Decision making: who’s in charge?

Decision-making can be regarded as a problem-solving activity terminated by a solution deemed to be optimal, or at least satisfactory. It is therefore a process which can be more or less rational or irrational and can be based on explicit or tacit knowledge and beliefs.

(Wikipedia)



A TENTATIVE ARCHITECTURE

Inside the brain

Cortex
- Posterior
- Motor / Premotor
- Prefrontal

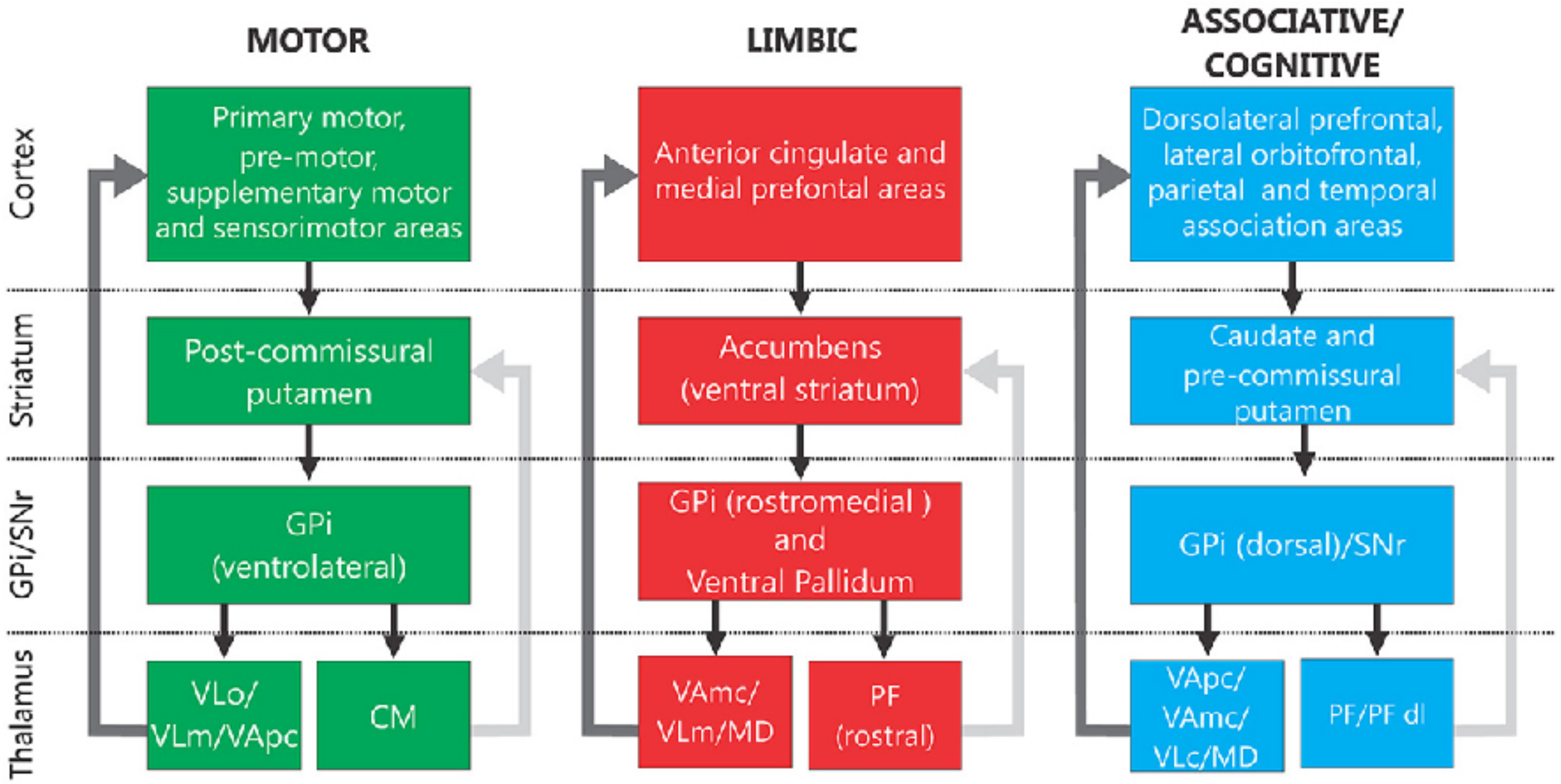
Thalamus

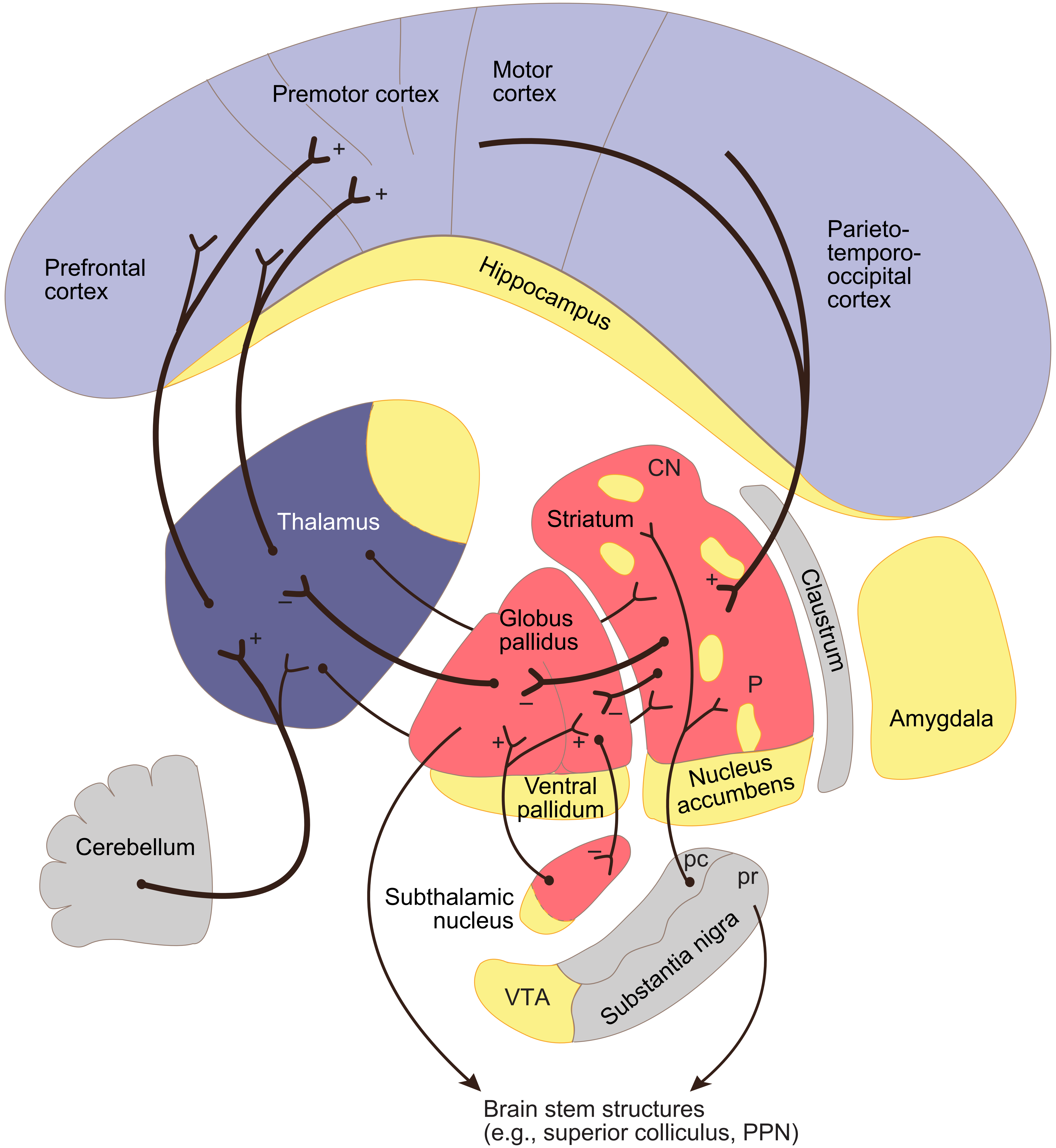
Striatum (STR)
- Caudate
- Putamen
- Nucleus Accumbens

Subthalamic Nucleus (STN)

Globus Pallidus
- Internal (GPi)
- External (GPe)

Subtantia Nigra
- pars Compacta (SNc)
- pars Reticulata (SNr)





A TENTATIVE SOLUTION

Decision making: who's in charge?

The executive decision maker
Brazil, Terry Gilliam, 1985

A SIMPLE MODEL

A binary choice

Let us consider a choice between an option **X** and an option **Y**

- We want the model to be simple
- We want the decision to be gradual
- We want the model to choose either option **X** or option **Y**
- We need a decision threshold

A SIMPLE MODEL

A dynamic system with 2 variables

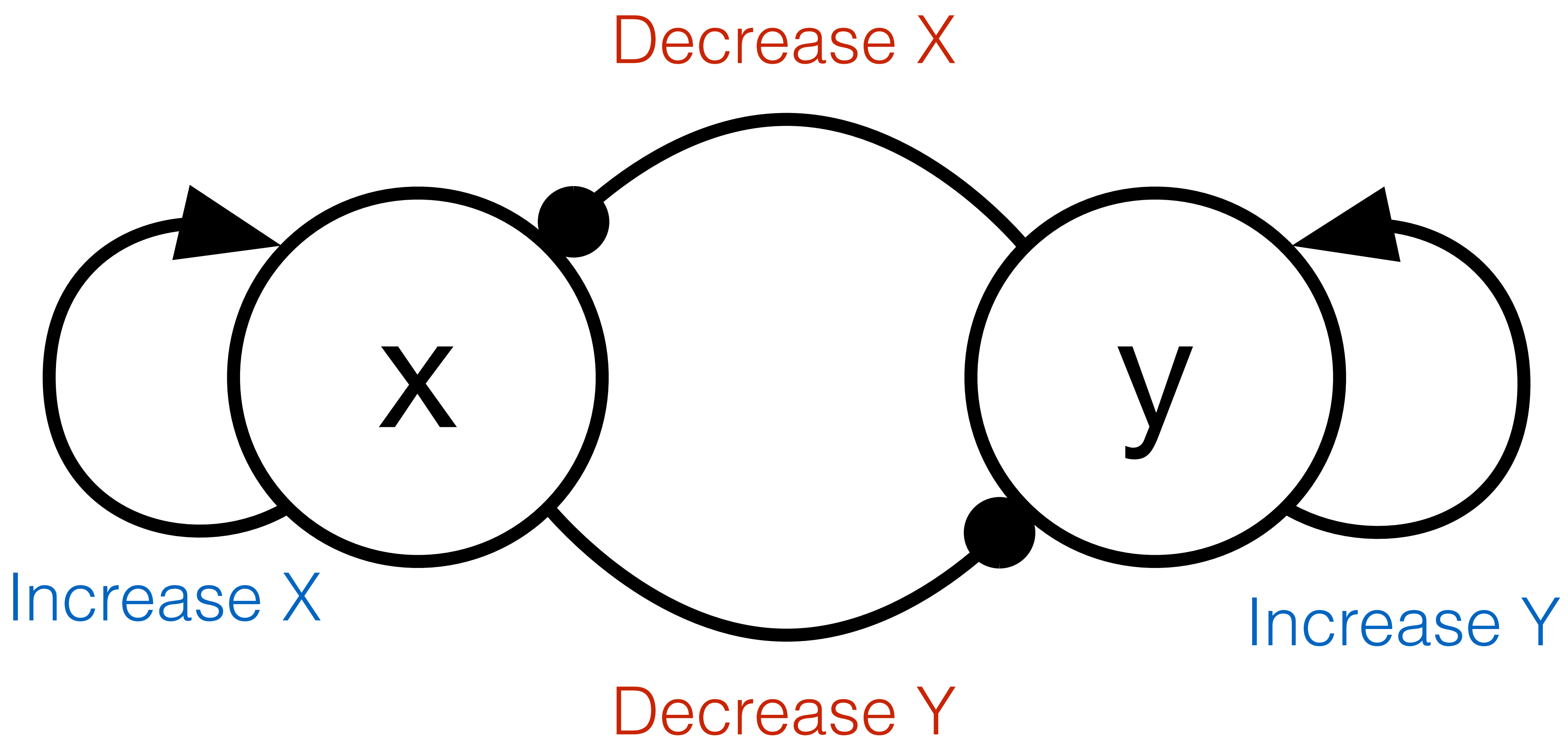
Let us consider a choice between an option **X** and an option **Y**

- We want the model to be simple
- We want the decision to be gradual
- We want the model to choose either option **X** or option **Y**
- We need a decision threshold

We'll consider a simple dynamical system of two variables

- **x** and **y** are two variables whose initial value is 0
- Each variable grows with time.
- Each variable influence the other variable such that
  - When **x** increases, it tends to make **y** to decrease
  - When **y** increases, it tends to make **x** to decrease
- If a variable reaches the value 1, a decision has been made

$$\dot{x} = \alpha(1-x) + (x-y)(1-x), x > 0$$
$$\dot{y} = \alpha(1-y) + (y-x)(1-y), y > 0$$



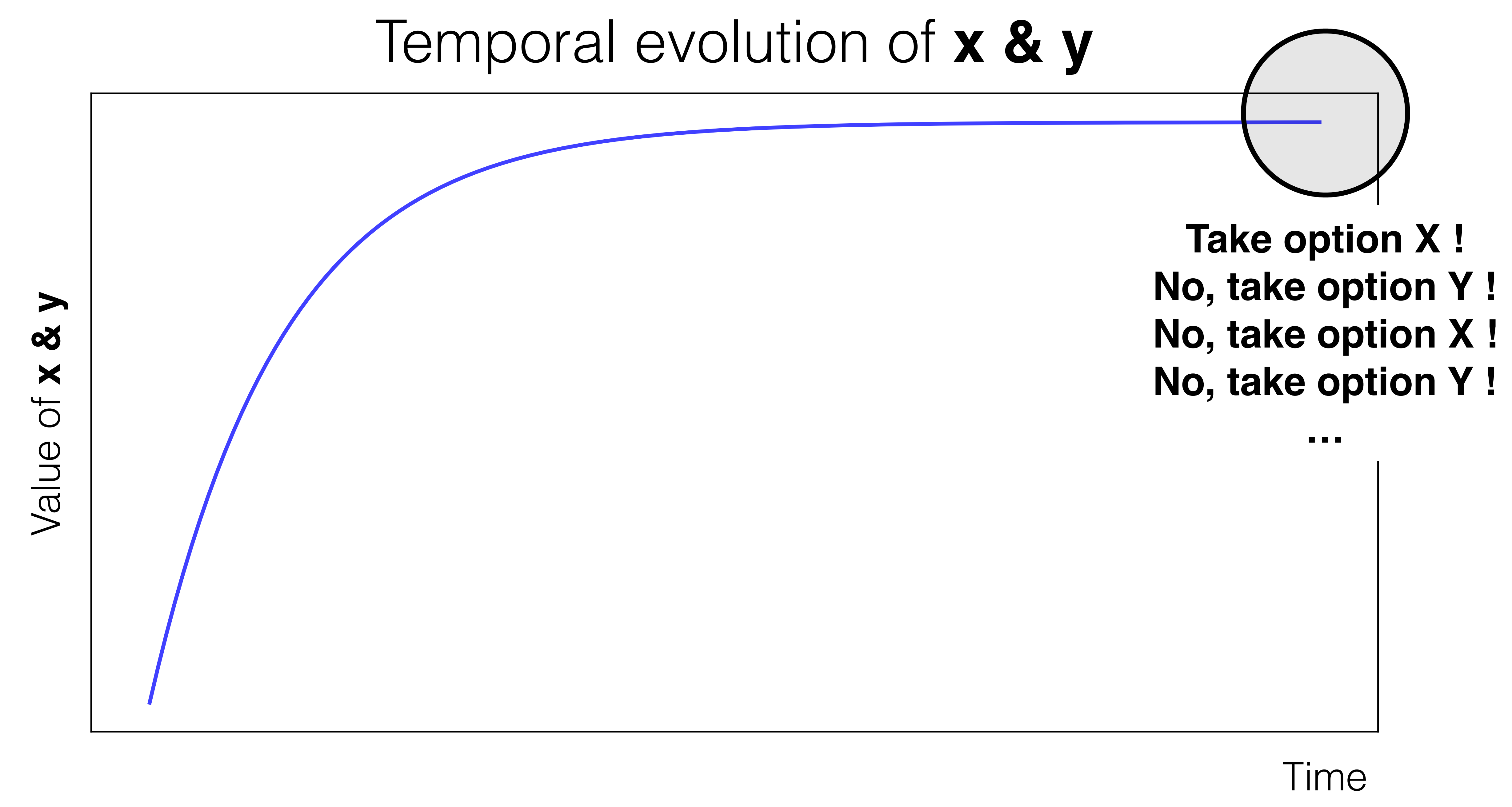
A SIMPLE MODEL
Taking a decision

We consider a simple dynamical system of two variables

- **x** and **y** are two variables whose initial value is 0
- Each variable grows with time.
- Each variable influence the other variable such that
  - When **x** increases, it tends to make **y** to decrease
  - When **y** increases, it tends to make **x** to decrease
- If a variable reaches the **value 1**, a decision has been made

When simulated, this system gives no decision…
→ We need to break the symmetry in the system, but how ?

$$\dot{x} = \alpha(1-x) + (x-y)(1-x), x > 0$$
$$\dot{y} = \alpha(1-y) + (y-x)(1-y), y > 0$$



Mersenne Twister to the rescue

There is no actual randomness in a computer.
This is the reason why we need to simulate it.

The Mersenne Twister is a pseudorandom number generator (PRNG) developed in 1997 by **Makoto Matsumoto** and **Takuji Nishimura**. It is the most widely used general-purpose PRNG. Its name derives from the fact that its very long period length is chosen to be a Mersenne prime ($2^{19937} - 1$).

Binary sequence examples:

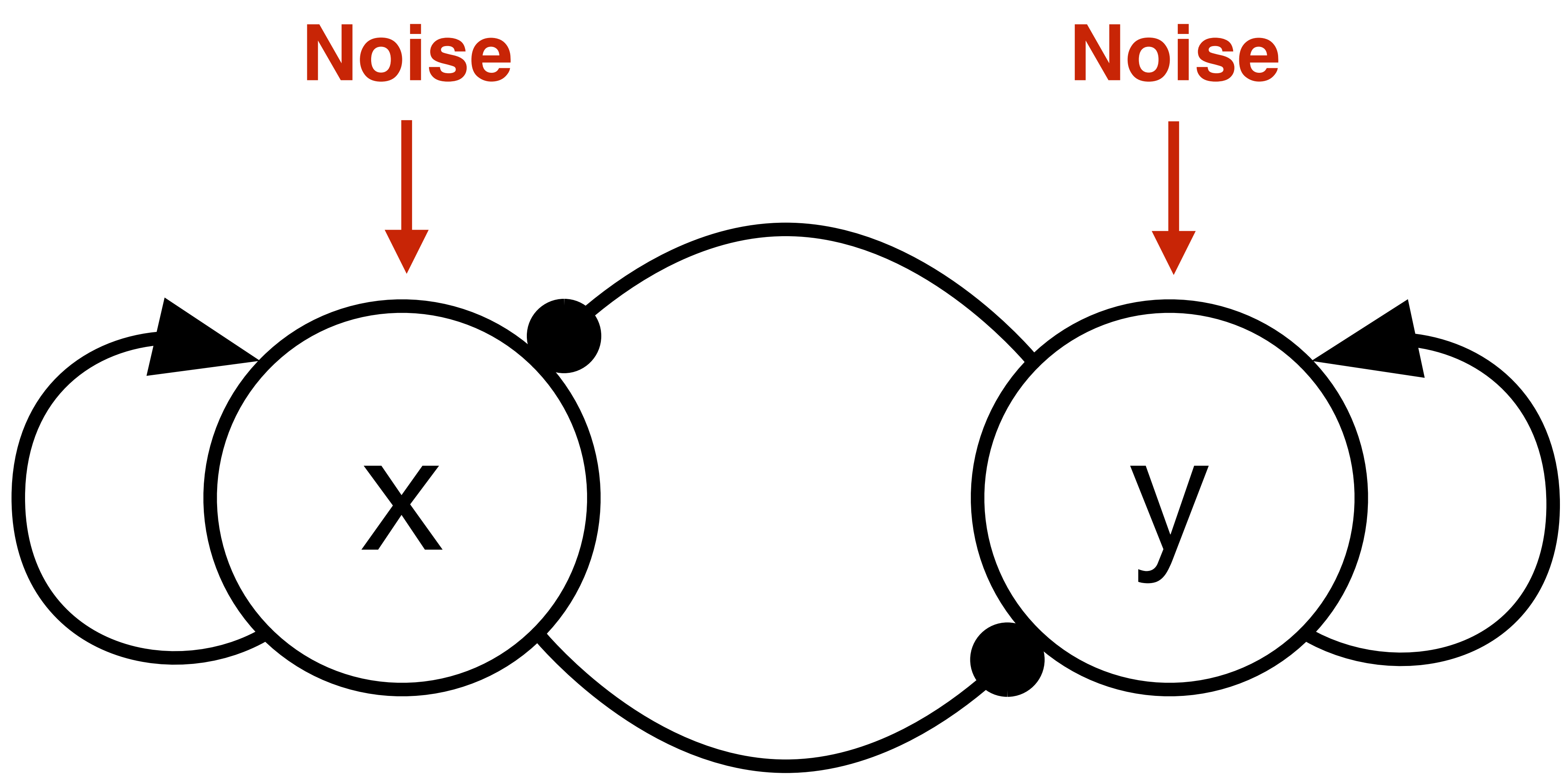
```
0,1,1,1,0,0,1,0,0,1,0,0,0,0,1,1,0,1
1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1  ←
1,0,0,1,0,1,0,0,1,1,0,1,1,1,1,1,1,0
1,1,1,0,0,0,0,0,1,1,1,0,0,0,1,0,0,1
```

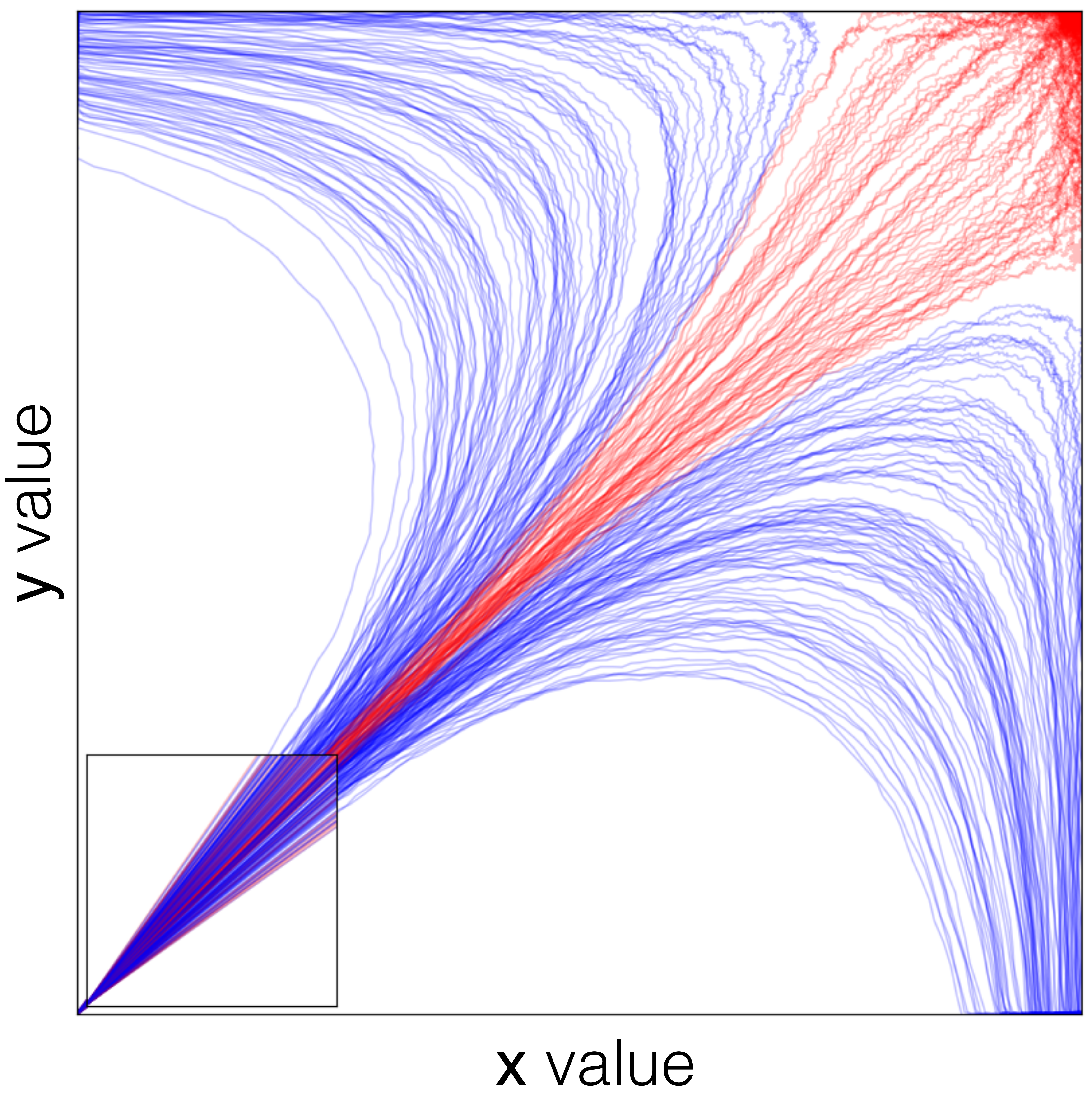
THIS ONE'S VALID AND AS MUCH PROBABLE AS ANY OTHER

A SIMPLE MODEL

Noisy decision

Noise from the sensors and the actuators
Environmental noise from the outside world
Numerical noise inside the model



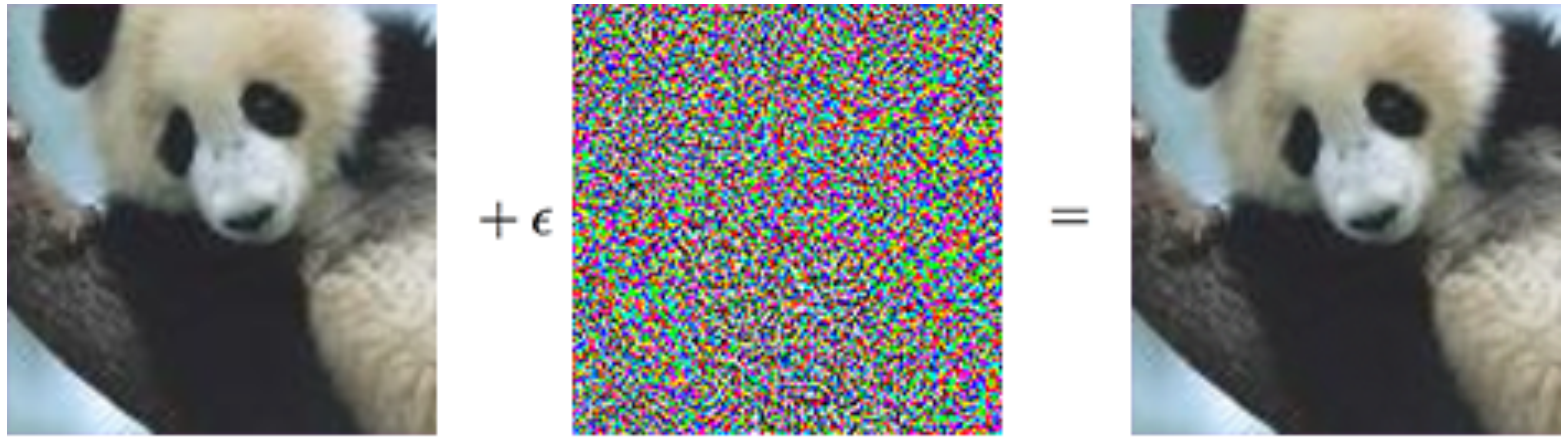


A SIMPLE MODEL

Noisy world

When there is no actual knowledge of what are objects, they're only recognized through a set of (learned) statistical features. Such statistical inferences can be tricked.

Perceptions and actions are noisy. A robot has to cope with it.



"panda"
57.7% confidence

"gibbon"
99.3% confidence

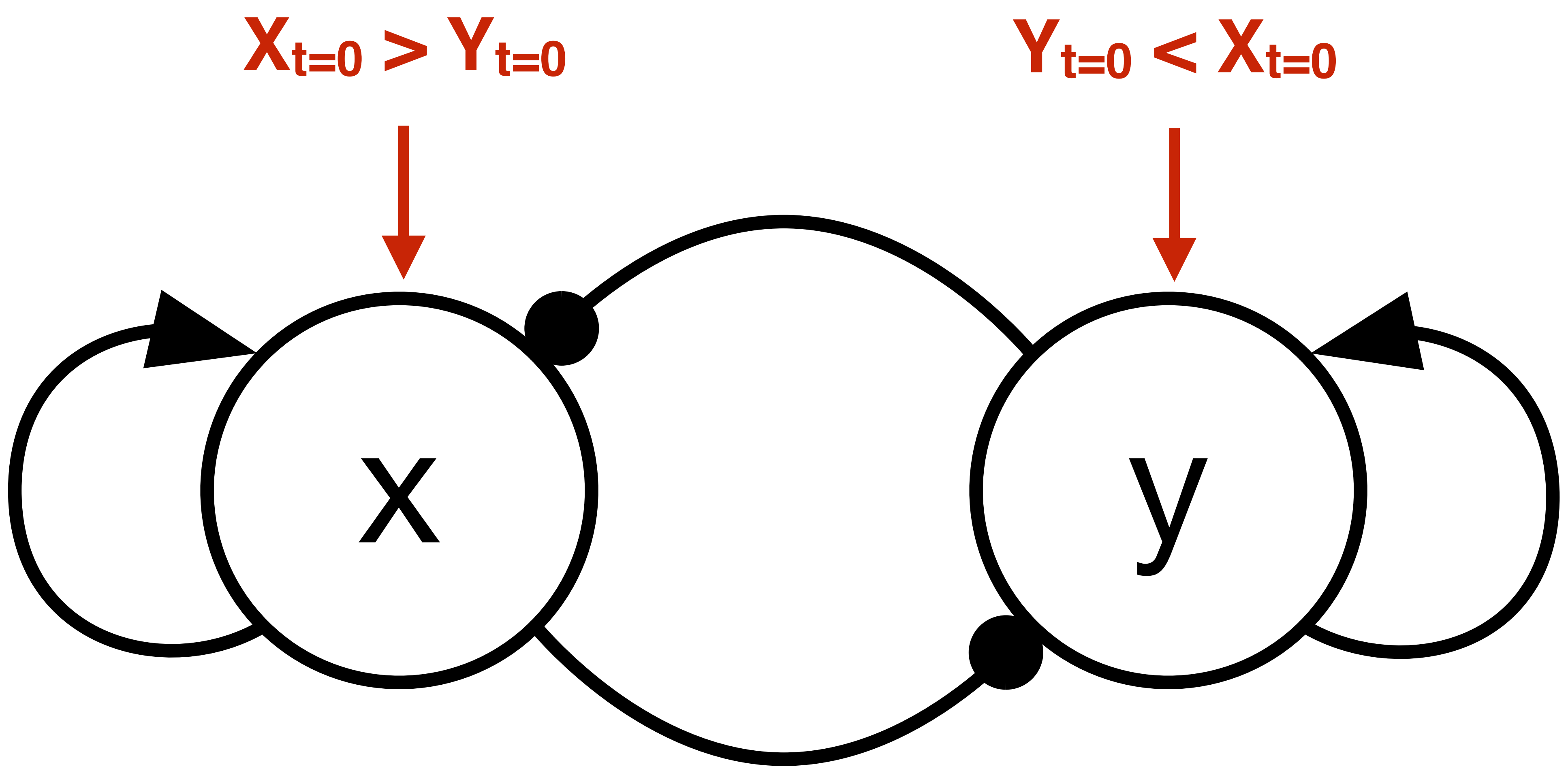
Reese Witherspoon

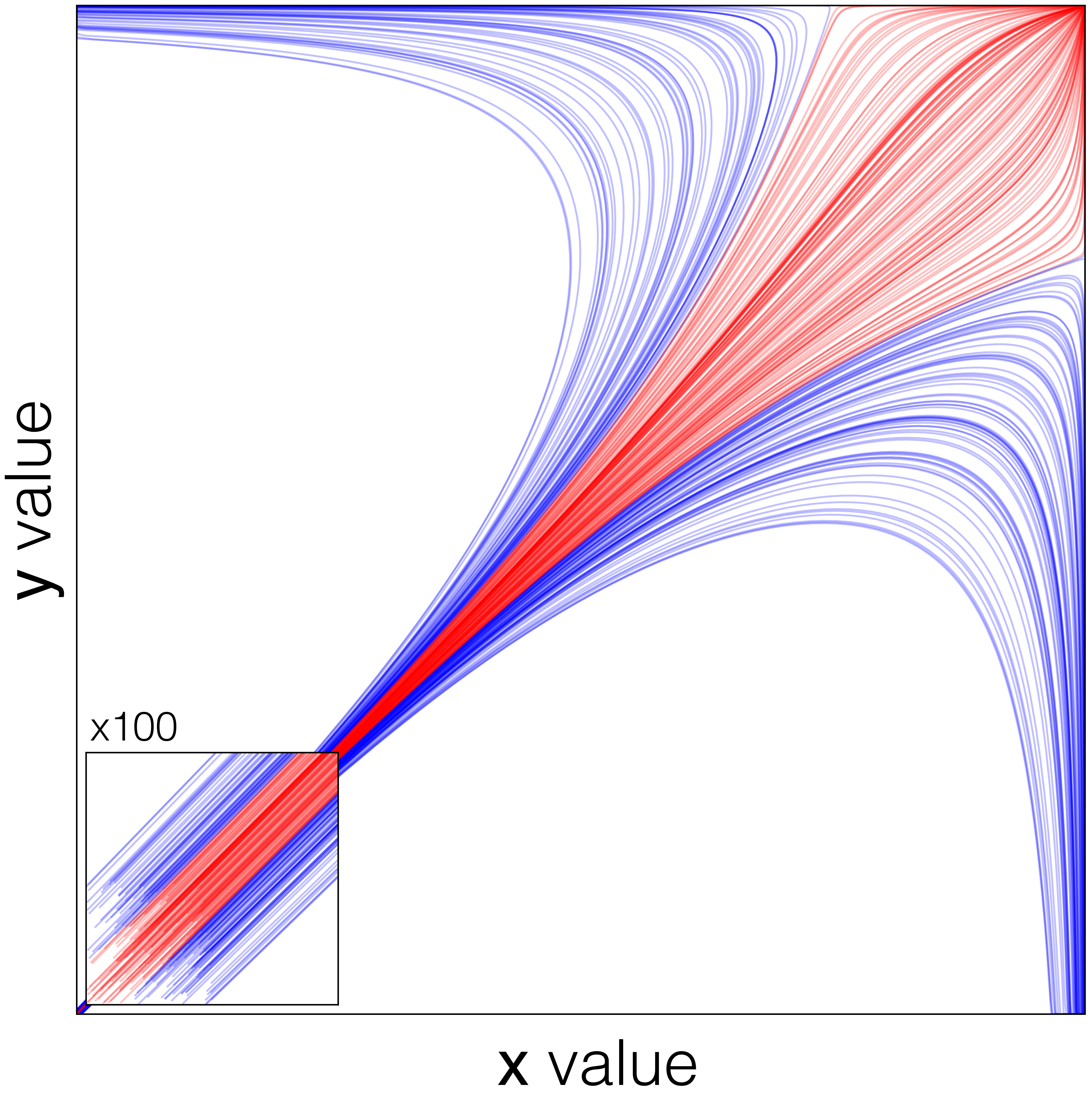
Reese Witherspoon
+ googles =

Russel Crowe

Initial conditions

Two similar situations are never actually the same.
Small initial differences can lead to great consequences.
Butterfly effect.



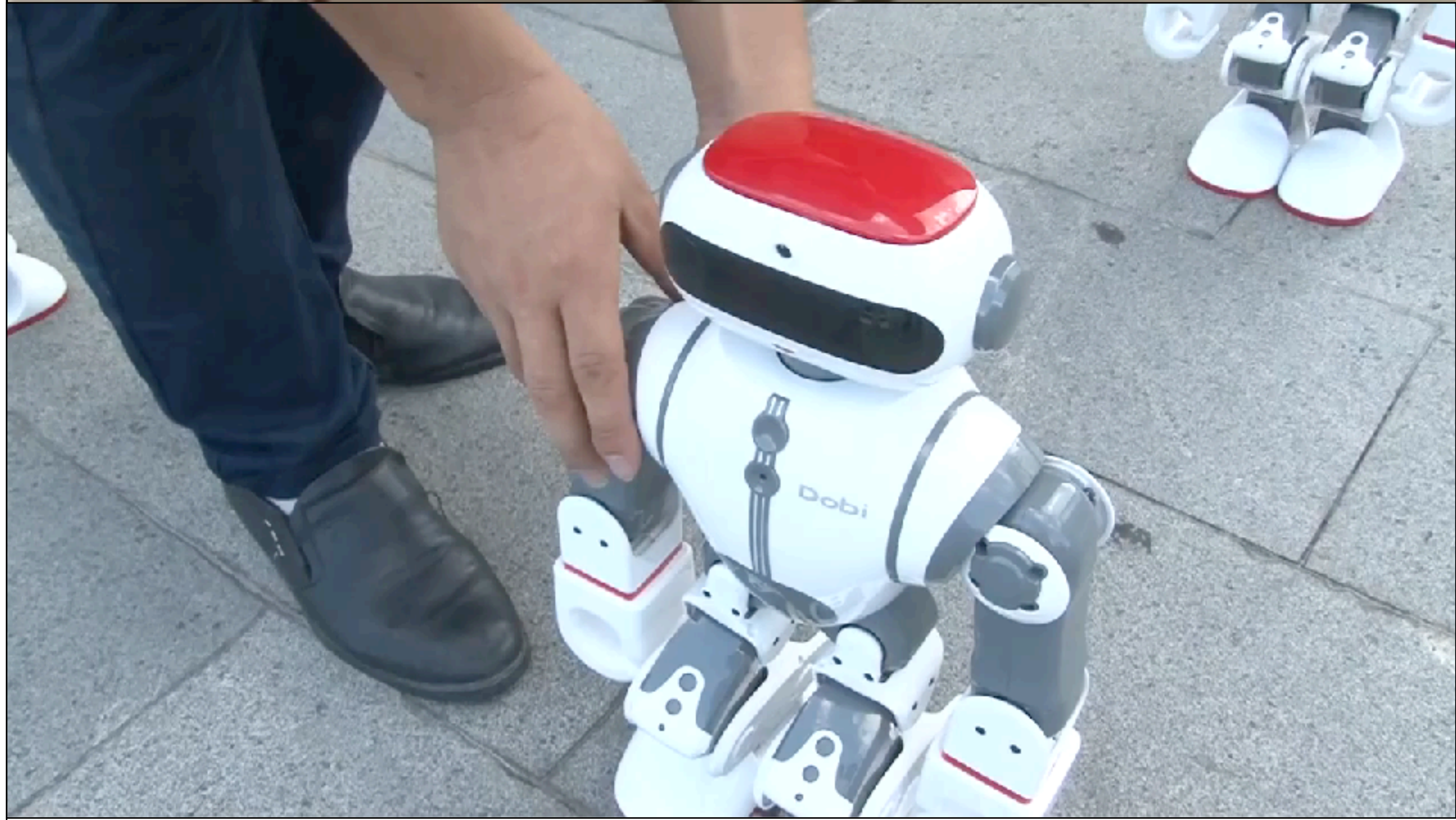


A SIMPLE MODEL

Initial conditions

The slightest difference between the model of the world and the actual world may have dramatic consequences.

The (dynamic) world is its own best model.



Learning to decide

After each decision (**X** or **Y**), a reward is given. The goal is to maximise the amount of reward through several choices:

Non probabilistic

X	**X**	X	X	X	**X**	X	**X**	**X**	X	X
Y	Y	**Y**	**Y**	**Y**	Y	**Y**	Y	Y	**Y**	**Y**
1	1	0	0	0	1	0	1	1	0	0

Probabilistic

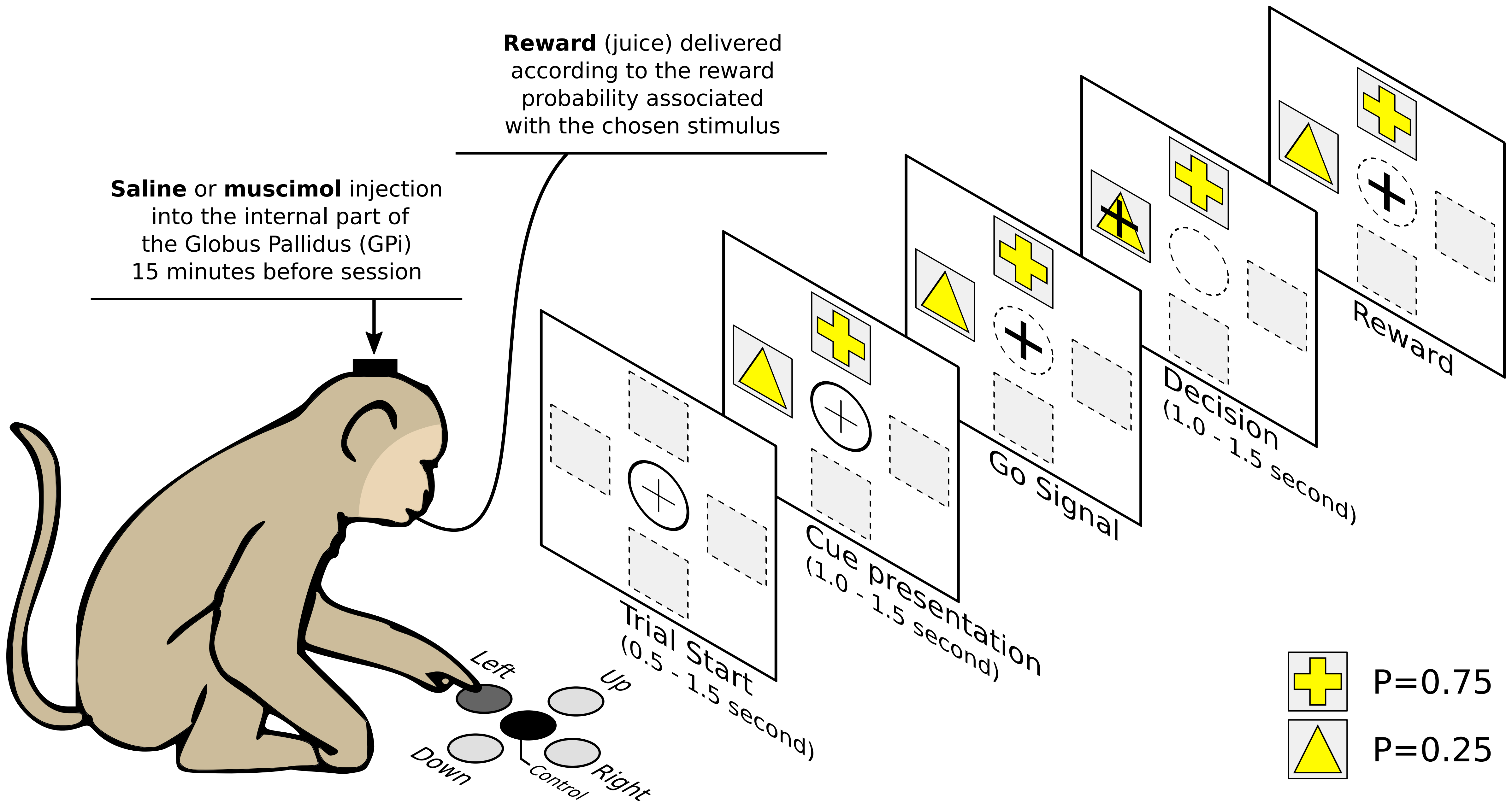
X	**X**	X	X	X	**X**	X	**X**	**X**	X	X
Y	Y	**Y**	**Y**	**Y**	Y	**Y**	Y	Y	**Y**	**Y**
1	0	1	1	0	0	1	0	1	1	0

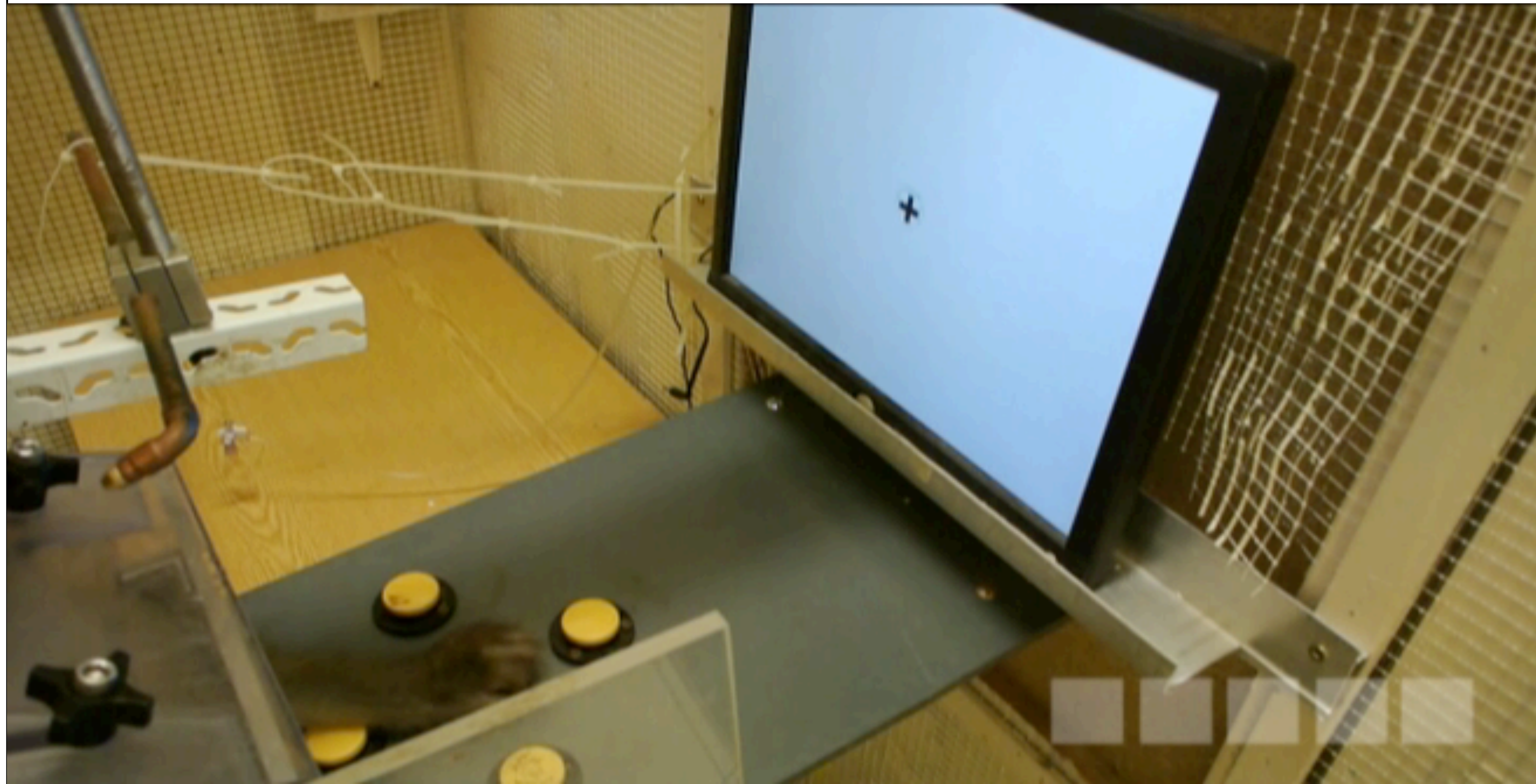
Monkeys can do it too

We designed a simple two-armed bandit task where two stimuli A and B are associated with different reward probability (respectively 0.25 and 0.75). The goal for the subject is to choose the stimulus associated with the highest reward probability, independently of its position.

X	**X**	X	X	X	**X**	X	**X**	**X**	X	X
Y	Y	**Y**	**Y**	**Y**	Y	**Y**	Y	Y	**Y**	**Y**
1	0	1	1	0	0	1	0	1	1	0

Exploration / exploitation dilemma

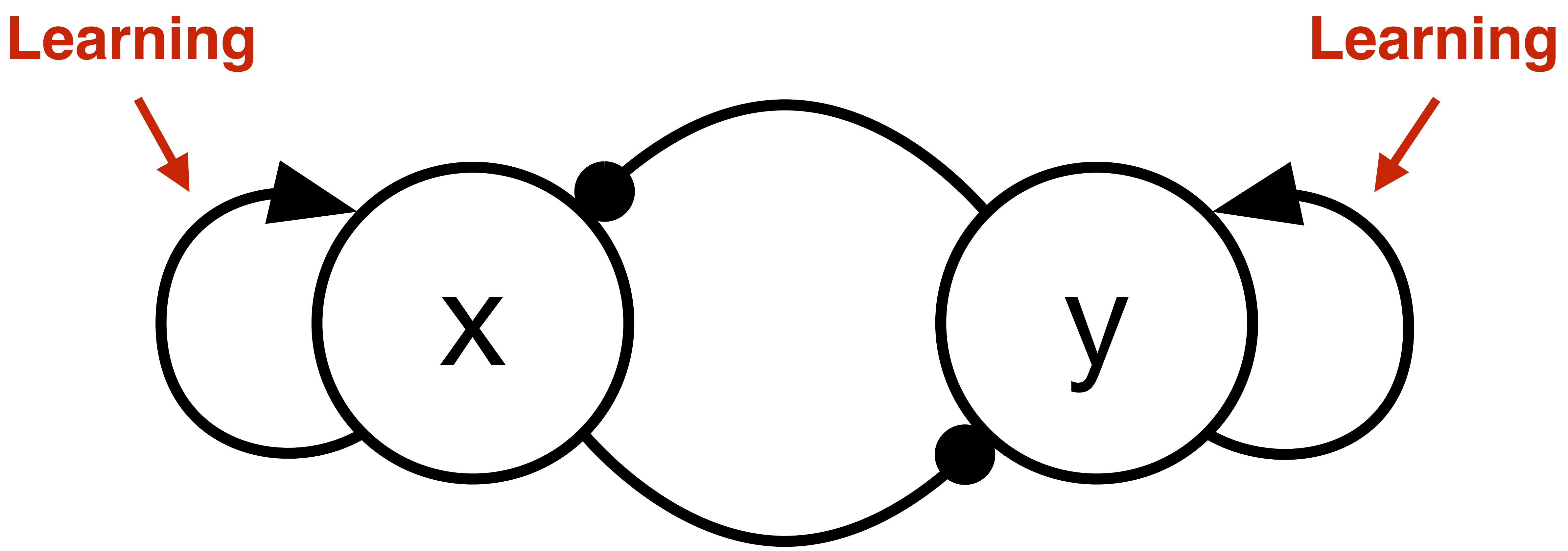


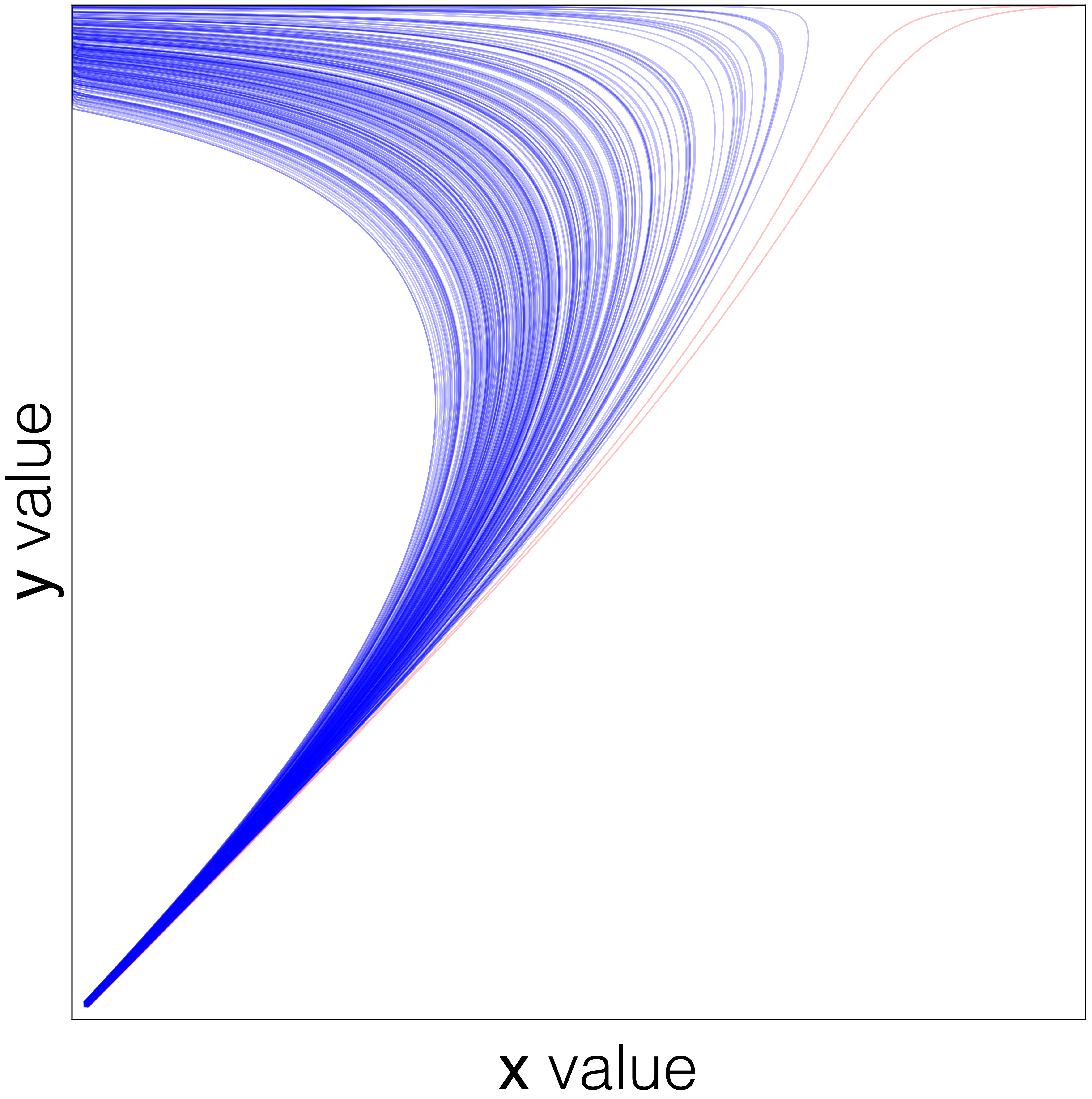


A SIMPLE MODEL

Learning to decide

- After a choice **X**, if a reward is received, the likeliness of choosing **X** next time is increased.
- After a choice **Y**, if a reward is received, the likeliness of choosing **Y** next time is increased.
- After a choice **X**, if no reward is received, the likeliness of choosing **X** next time is decreased.
- After a choice **Y**, if no reward is received, the likeliness of choosing **Y** next time is decreased.





A SIMPLE MODEL

Biased dataset

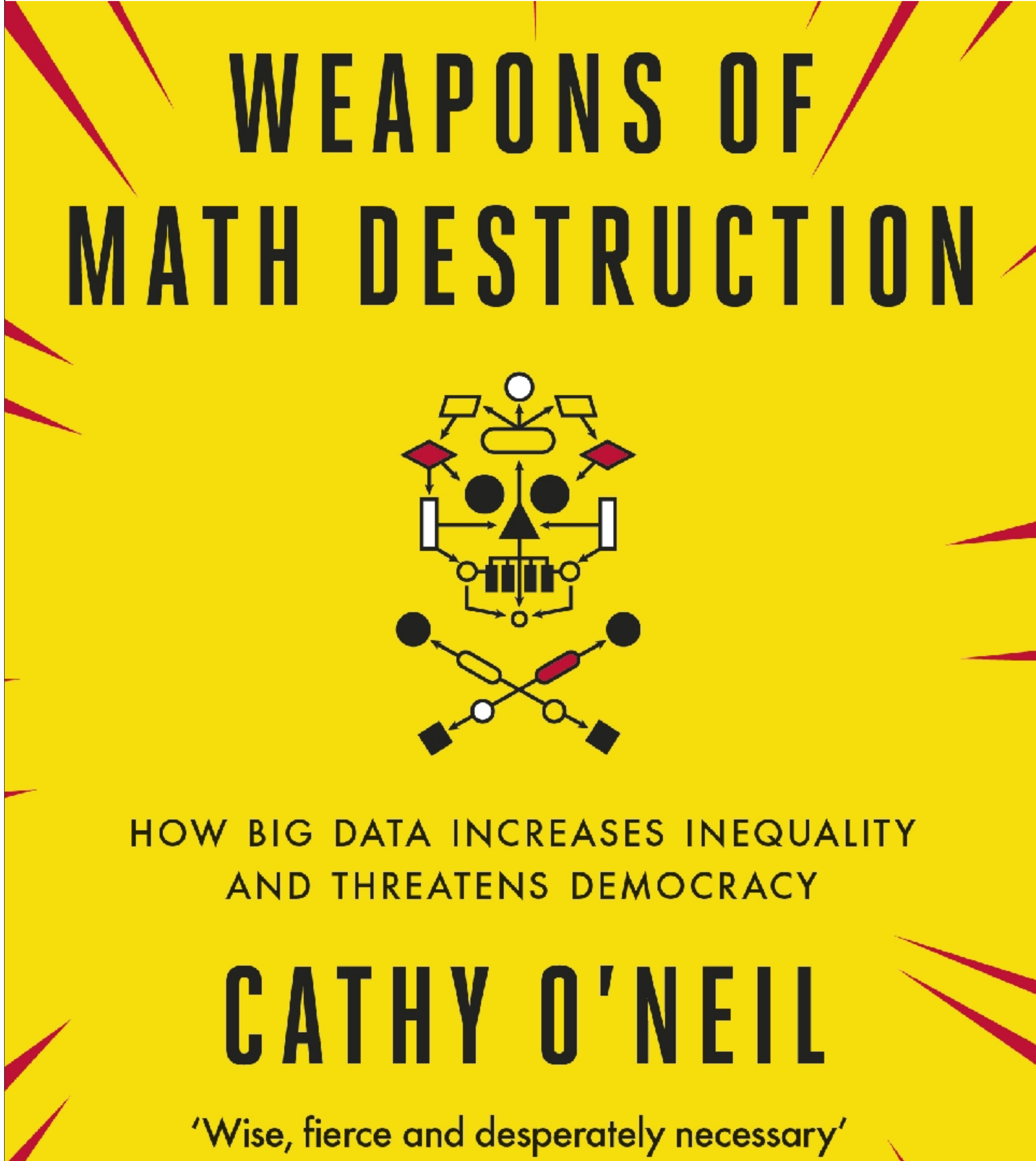
If a bias exists in the dataset, this bias will most likely be learned by the model.

Dataset 1 (Reward: X=75%, Y=25%)

X	**X**	X	X	X	**X**	X	**X**	**X**	X	X
Y	Y	**Y**	**Y**	**Y**	Y	**Y**	Y	Y	**Y**	**Y**
1	0	0	0	1	1	1	0	1	0	0

Dataset 2 (Reward X=75%, Y=25%, biased)

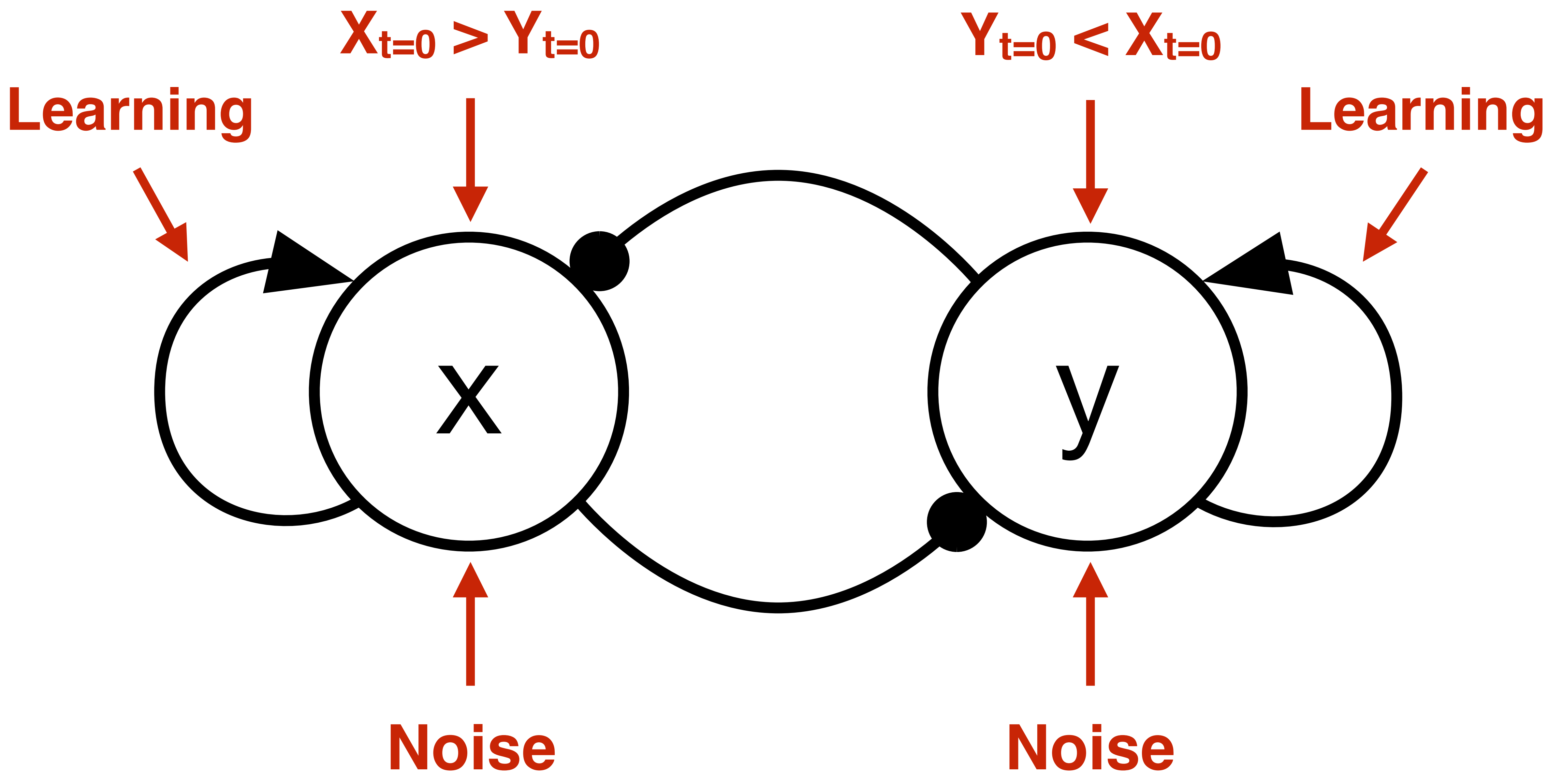
X	**X**	X	X	X	**X**	X	**X**	**X**	X	X
Y	Y	**Y**	**Y**	**Y**	Y	**Y**	Y	Y	**Y**	**Y**
0	0	1	1	0	0	1	0	1	1	1

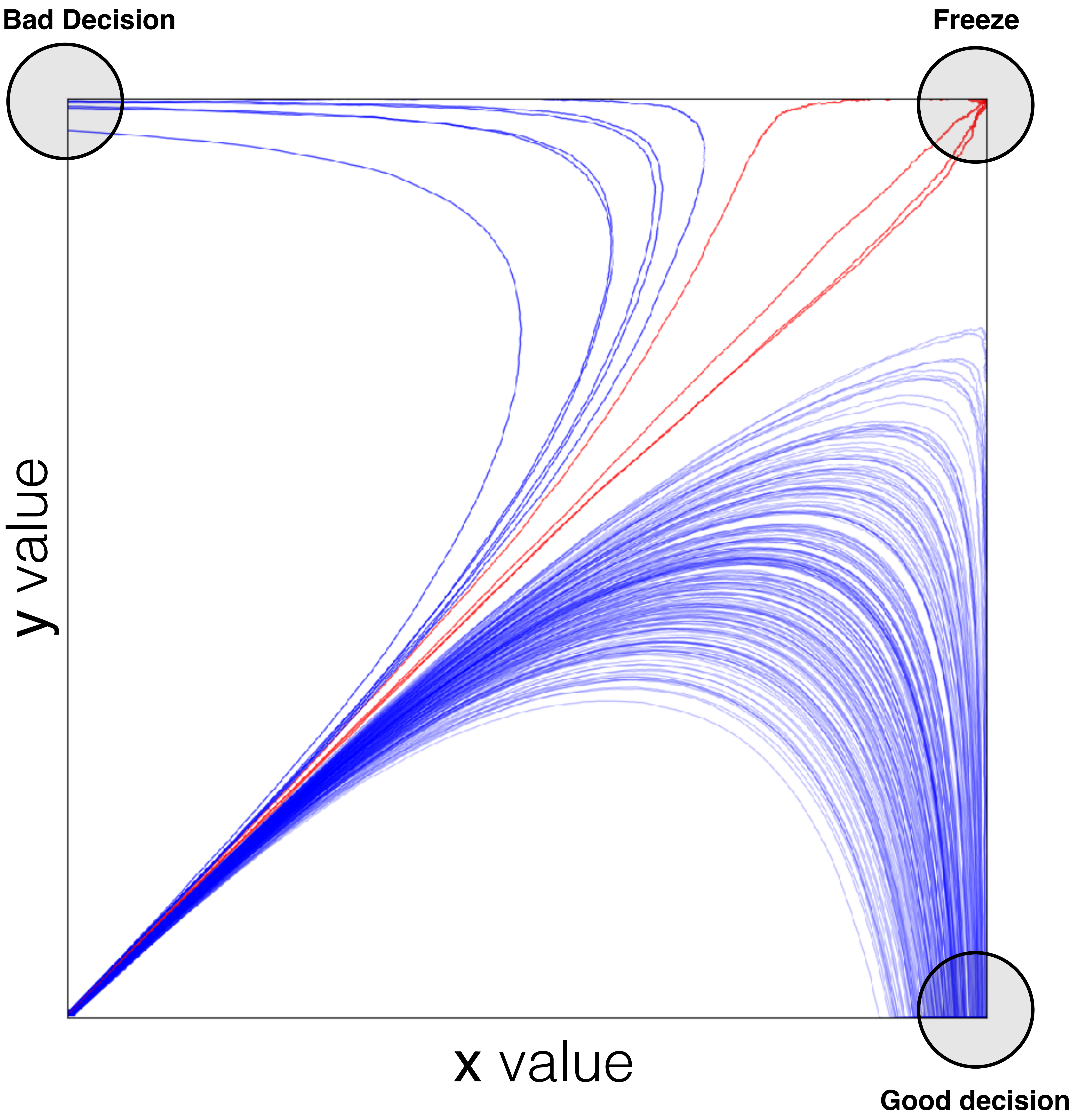


A SIMPLE MODEL

Putting all together

When observing successive decisions, we can see a clear tendency to prefer **X** over **Y**, but from time to time, the model will either choose **Y** or make no choice at all.





Conclusion

A decision in a model (robot) results from an equilibrium between several interlinked factors at different levels:

- Noise (the origin of everything…)
- Data (biases in the data will most likely be learned)
- Model (the choice of the model governs the overall behavior)
- Learning (Full learning autonomy (or not))
- Implementation (the code might be plainly wrong)
- Environment (biases may exist in the software stack and/or OS)

Ultimately, it is quite complicated (or impossible) to explain why an individual decision has been taken. It is much easier to explain the overall behavior in term of average performances.

But this might be hardly satisfactory to explain an isolated weird or bad decision (Tesla, 2016 "Oops, sorry, you're dead", Microsoft 2016, Tay chatbot "incident")

We tend to mostly notice weird events.
The higher the expectation, the higher the disappointment.

DECISION MAKING: WHO’S IN CHARGE?